RESEARCH   **Open Access**

Investigation of inter-rater and test-retest reliability of Y balance test in college students with flexible flatfoot

Yalin Zheng[1], Renzhi Feng[1], Weiyin Hu[1] and Peng Huang[1*]

Abstract

Background The Lower Quarter Y Balance Test (YBT-LQ) has been widely used to assess dynamic balance in various populations. Dynamic balance in flexible flatfoot populations is one of the risk factors for lower extremity injuries, especially in college populations in which more exercise is advocated. However, no study has demonstrated the reliability of the YBT-LQ in a college student flexible flatfoot population.

Methods A cross-sectional observational study. 30 college students with flexible flatfoot were recruited from Beijing Sports University. They have been thrice assessed for the maximal reach distance of YBT under the support of the lower limb on the flatfoot side. Test and retest were performed with an interval of 14 days. The outcome measures using the composite score and normalized maximal reach distances in three directions (anterior, posteromedial, and posterolateral). The relative reliability was reported as the Intraclass Correlation Coefficient (ICC). Minimal Detectable Change (MDC), Smallest worthwhile change (SWC), and Standard Error of Measurement (SEM) were used to report the absolute reliability.

Results For inter-rater reliability, the ICC values for all directions ranged from 0.84 to 0.92, SEM values ranged from 2.01 to 3.10%, SWC values ranged from 3.67 to 5.12%, and MDC95% values ranged from 5.58 to 8.60%. For test-retest reliability, the ICC values for all directions ranged from 0.81 to 0.92, SEM values ranged from 1.80 to 2.97%, SWC values ranged from 3.75 to 5.61%, and MDC95% values ranged from 4.98 to 8.24%.

Conclusions The YBT-LQ has "good" to "excellent" inter-rater and test-retest reliability. It appears to be a reliable assessment to use with college students with flexible flatfoot.

Trial registration This trial was prospectively registered at the Chinese Clinical Trial Registry with the ID number ChiCTR2300075906 on 19/09/2023.

Keywords Flexible flatfoot, College Students, Y Balance Test Lower quarter, Reliability

*Correspondence:
Peng Huang
huangp@bsu.edu.cn
[1]School of Sports Medicine and Rehabilitation, Beijing Sports University,
No.48 Xinxi Road, Haidian District, 100084 Beijing, China

Background

Flexible flatfoot (FFF) is a common musculoskeletal pathology characterized by a collapsed medial longitudinal arch, forefoot abduction and pronation, and hindfoot valgus in the weight-bearing position, which recovers to normal in the non-weight-bearing [1]. Previous studies demonstrated the prevalence of flatfoot between 19.0%~26.5% in various ages and populations [2]. Moreover, earlier research indicates that flatfoot can impair the mobility of the medial longitudinal arch and increase foot stress, which can result in posterior tibial muscle dysfunction and patellofemoral instability [3], thus affecting the individual's proprioception, balance, and sports performance [4, 5]. Abnormal structural alterations in the foot have been recognized as a risk factor for lower limb sports injuries, specifically in individuals with flexible flatfoot. The risk of knee injuries, soft tissue injuries, and medial tibial stress syndrome (MTSS) is raised as a result of these changes [6].

Dynamic balance, which refers to an individual's capacity to maintain a stable center of gravity while exercising, is recognized as one of the risk factors for lower limb sports injuries. Moreover, it has been suggested that it may be a predictor of injury risk in athletes [7, 8]. The YBT-LQ, derived from the Star Excursion Balance Test (SEBT), is an economical and efficient assessment of dynamic balance. It has gained popularity for its application in injury prevention and screening among athletes [9, 10], particularly for lower limb injuries like ankle and knee injuries [11], Additionally, previous studies have shown the potential of this test to detect balance deficits in patients with low back pain [12]. YBT-LQ not only enhances the reproducibility of SEBT but also addresses common errors in SEBT measurements and standardizes the test procedure. The advantage of YBT-LQ over SEBT lies in its standardized protocol, which enables researchers and clinicians to compare results [13]. Furthermore, it has been observed to exhibit high interrater reliability (ICC0.73~1.00) and test-retest reliability (ICC0.68~0.94) in healthy populations [14].

The physical and mental health of college students tends to continuously decline worldwide [15]. Thus, college students should be encouraged to participate in more physical exercise and enhance their physical and mental fitness, to lay a foundation for their healthy growth and lifelong development [16]. The sport demands unilateral balance and dynamic neuromuscular control. Changes in foot structure can alter contact area, joint motion, and muscle activation strategies in the flexible flatfoot population, which can have a negative impact on balance and increase the risk of lower limb sports injuries [17].

The Y Balance Test (YBT) is a low-cost clinical measure of dynamic balance that simulates the demands of exercise requiring unilateral balance, and it has also been used as a clinical outcome measure to measure functional improvement and guide activity progression after injury [18–20]. The YBT is frequently used as an outcome measure of dynamic balance ability in studies related to dynamic balance in patients with flatfoot and provides a better response to detecting deficits in their dynamic balance [4, 21]. However, there is still a gap in the research on the reliability of the Y balance test in the flatfoot population.

However, the current reliability studies of the YBT-LQ have focused primarily on healthy populations and athletes [22, 23]. Although the YBT is currently considered a reliable tool for measuring dynamic balance and predicting injury risk in healthy populations and athletes, there are differences between these populations and college student flatfoot populations.This may limit the generalisability of these findings to the flexible flatfoot college students. The utility of the YBT-LQ to assess dynamic balance in a college student flatfoot population is currently unknown.

Therefore, this cross-sectional observational study aimed to further investigate the reliability of the YBT-LQ in a flexible flatfoot population of college students. We assumed that the relative reliability reported as ICC would be good, corresponding to an ICC>0.75 indexed by Koo et al. [24]. The absolute reliability reported in terms of SEM and MDC would be on a similar level as found by Shaffer et al. [25]. The value of SWC would be larger than SEM, and YBT can detect the smallest worthwhile changes.

Methods

Study design

This cross-sectional observational study was conducted between September 10 and October 20, 2023, at Beijing Sports University, Beijing, China. The study adhered to STROBE guidelines.

Sample size calculation

A sample size of at least 30 participants was calculated in advance by PASS 2021(NCSS LLC., Kaysville, U.T., USA) to allow for an inter-rater test-retest reliability study that would result in an ICC of at least 0.75, with an expected 0.9 and beta of 0.8. alpha was 0.05, and the dropout rate was 10%.

Participants

30 Participants were recruited from the Beijing Sports University through announcements and personal referrals. Participants were eligible for inclusion if they were 17~25 years college students; had navicular drop test>10 mm; had no history of lower limb injury within 6 months; had no history of surgery on hip, knee, and ankle joints; and had no serious medical conditions.

Participants were excluded if they had vestibular dysfunction or other diseases that affected balance; taking unknown drugs; rheumatoid, neurological, or other causes that affected the hip and knee joints muscles or function.

Written informed consent was obtained from the participants before the start of the study. The Sports Science Experiment Ethics Committee of Beijing Sport University approved the study protocol (2,023,154 H). This trial was registered at the Chinese Clinical Trial Registry (ChiCTR2300075906) on 19/09/2023.

Testing procedures

The Guidelines for Reporting Reliability and Agreement Studies were used to implement and report this interrater and test-retest reliability study [26]. To examine the interrater and test-retest reliability of the YBT, the extent of agreement and reproducibility were calculated between the measurements by two different raters who were fully experienced in the use of YBT. In the first trial, the participant would perform 6 practice sessions to eliminate the learning effect before the formal YBT-LQ and perform the test by the rater within 20 min of the practice session. Subsequently, after a 20-minute break, the test was administered by another rater. A 14-day interval was set between the test and the retest to prevent actual changes in dynamic balance ability and memory effects.

All raters were pre-trained and proficient in the testing process. Hiding test data immediately after completion of the test so that the rater did not have access to each subject's previous score to minimize bias. All test data would be double-entered into the computer to ensure proper.

Navicular drop test

The navicular drop test (NDT) was performed to measure the change in height of the medial longitudinal arch of the foot. The participant seated in a chair with the knee flexed to 90° and the second toe aligned with the knee so that the talonavicular joint was in a neutral position. The height between the ground and the navicular tuberosity was measured and noted by the evaluator using a vernier caliper. Then, the participant stood up and measured the height again. The difference between the height of the navicular tubercle in the non-weight-bearing and weight-bearing positions of the foot was calculated. Repeated the test 3 times on both the right and left sides and take the average value. Navicular drop exceeding 10 mm was diagnostic of a flexible flatfoot [27].

Leg length measure

Participants were instructed to lie in a supine position. They lifted their hips and returned to the starting position. To make sure the pelvis was in alignment, the participant were passively straightened the legs. The participants' flexible flatfoot side leg length was measured centimeters from the anterior superior iliac spine (ASIS) to the most distal part of the medial malleolus using a cloth tape. This measurement was used to normalize the maximal reach distance. The leg length measurement was carried out before the first trial.

Y balance test lower quarter

The test was carried out by the standardized YBT methodology advised by Plisky et al. [13]. The trial protocol was conducted on the side leg with the flexible flatfoot, which was the only difference from the YBT-LQ. The YBT-LQ was completed using the Y Balance Test Kit. Before the test, participants were shown the YBT demonstration video and provided detailed instructions on how to take it. According to the previous research protocol [28–30], the practice phase consisted of six trials in three different directions in order to minimize the learning effect. Participants performed the YBT-LQ at least 3 times and up to 6 times in each direction after the practice phase. All tests were conducted with participants standing barefoot on the pedals, with their single lower limb supported by the flatfoot side. The distal end of the second toe of the supporting leg was located behind the red indicator line, and the contralateral lower limb was extended in three directions, pushing the rectangular indicator block as far as possible with the tip of the foot. Then, recorded the reading of the proximal end of the rectangular block to the nearest 0.5 cm. The outcome of that test was deemed invalid if participants were (1) unable to controllably return to the starting position, (2) accelerate the rectangular block with the outstretched foot to move it farther, (3) touch the ground with their forefoot, (4) or contact the top of the rectangular block for support. The maximal reach distance in each direction was recorded for the 3 valid tests.

The normalized maximal reach distance per reach direction was calculated as follows (Eq. 1) and used as an outcome measure. Additionally, Filipa et al. [31] supplied a formula (Eq. 2) that was used to determine the normalized composite score (CS).

$$
\begin{aligned}
&Normalized\ maximal\ reach\ distance\,(\%\ leg\ length\,[\,LL\,]) = \\
&(absolute\ maximal\ reach\ distance\,[cm])\,/LL\,[cm])\times\ 100\%
\end{aligned} \tag{1}
$$

$$
\begin{aligned}
&Composite\ score\,(\%LL) = \\
&\{(AT + PM + PL)\,/\,(leg\ length \times 3)\}\ \times 100\%
\end{aligned} \tag{2}
$$

Statistical analysis

Reliability refers to the consistency of a test or measurement [32]. Calculating the ICC can demonstrate relative reliability, which is the degree to which individuals maintain their location in a sample with repeated

Table 1 Characteristics for participants performing the Y Balance Test

	Mean ± SD[95%CI]	Minimum;Maximum
Sex, n (%)	female 16(53.3%), male14(46.7%)	-
Age(years)	21.53 ± 0.32[20.88,22.18]	18;24
Height(cm)	170.47 ± 8.75[167.19,173.73]	156;186
Mass(kg)	64.65 ± 12.21[60.09,69.21]	44;87
BMI (kg/m²)	22.09 ± 2.75[21.06,23.11]	15.59;25.31
Length(cm)	85.04 ± 5.02[83.16,86.91]	75.3;92.2
Flexible flatfoot, n (%)	Left 6(20%), Right24(80%)	-

Values are presented with mean±standard deviation (SD) and 95% confidence intervals [95% CI].BMI=body mass index

measurements [33]. The maximal reach distance in each direction and composite score were calculated as absolute values and values normalized to leg length. The relative reliability of the maximal reach distance in each direction and of the composite score was analyzed by calculating ICC. The inter-rater reliability was determined using the ICC(2,1) and the test-retest reliability was determined using the ICC(3,1).ICC< 0.50 was considered "poor", 0.50<ICC< 0.75 was considered "moderate", 0.75<ICC< 0.9 was considered "good", and ICC≥0.90 was considered "excellent" [24].

The absolute reliability of the data was assessed using the standard error of measurement (SEM) that estimates the amount of error related to the measurement (Eq. 3). Moreover, the minimal detectable change (MDC95%) and the smallest worthwhile change(SWC) were calculated using the formula (Eq. 4 and Eq. 5) to ensure the differences between test and re-test measurements were real and outside the error range [24, 34]. It is worth mentioning that the study was conducted on a population of college students with flexible flatfoot, and there were no strict inclusion requirements for the fitness status of the participants, so the SWC calculated by 0.6 multiplied by SD is relevant for both high and low fitness level participants [34]. For all analyses, the Statistical Package for the Social Sciences (SPSS version 27.0) was used (IBM Corp., Armonk, N.Y., USA).

$$SEM = SD * \sqrt{1 - \text{ICC}} \qquad (3)$$

$$MDC95\% = 1.96 * \sqrt{2} * SEM \qquad (4)$$

$$SWC = 0.6 * SD \qquad (5)$$

Bland-Altman plot is a simple and intuitive graphical method of responding to data agreement. The 95% limits of agreement (LOA) was defined as two standard deviations above and below the mean of the difference scores. Plotting the difference between the YBT-LQ and comparison test values versus the mean of the YBT-LQ and

Table 2 The absolute and normalized maximal reach distance at test and retest in the three test directions (anterior, posteromedial, and posterolateral) and the composite score

Direction	Absolute Maximal Reach Distance(cm)Mean ± SD [95%CI]	Normalized Maximal Reach Distance (%) Mean ± SD[95%CI]
Rater A		
AT	58.65 ± 5.16[56.72,60.58]	69.11 ± 6.25[66.77,71.44]
PM	97.6 ± 9.99[93.87,101.33]	114.74 ± 9.15[111.33,118.16]
PL	94.32 ± 8.83[91.02,97.61]	110.97 ± 8.83[107.67,114.26]
CS	250.57 ± 20.16[243.04,258.09]	98.27 ± 6.37[95.89,100.64]
Rater B		
AT	57.80 ± 5.55[55.73,59.87]	68.14 ± 7.10[65.49,70.49]
PM	96.8 ± 8.70[93.55,100.05]	113.85 ± 7.95[110.88,116.82]
PL	94.08 ± 8.20[91.02,97.14]	110.73 ± 8.38[107.60,113.86]
CS	248.68 ± 17.94[241.99,255.38]	97.57 ± 5.95[95.35,99.79]
Retest		
AT	58.15 ± 5.17[56.22,60.08]	68.54 ± 6.43[66.14,70.94]
PM	97.07 ± 9.89[93.37,100.76]	114.16 ± 9.71[110.54,117.79]
PL	93.32 ± 8.65[90.09,96.55]	109.83 ± 9.03[106.45,113.20]
CS	248.53 ± 18.62[241.58,255.49]	97.51 ± 6.20[95.19,99.82]

Values are presented with mean±standard deviation (SD) and 95% confidence intervals [95% CI]

Normalized reach distance was calculated as (absolute reach distance/leg length) *100%

AT=anterior; PM=posteromedial; PL=posterolateral CS=composite score

comparison test scores yielded Bland-Altman plots with 95% limits of agreement, which were used to visually display measurement errors against true values. Statistical analyses were performed using MedCalc, version 19.4 for Windows (MedCalc Software, Ostend, Belgium).

Results

30 college students(16 females, 14 males;21.53±0.32years) participated in this study. The characteristics of the participants are shown in Table 1. No adverse events occurred during this study. There was no missing data in the current study.

The absolute and normalized maximal reach distance at the test for two raters and retest in the three test directions (anterior, posteromedial, and posterolateral) and the composite score are present in Table 2.

For inter-rater reliability, the ICC values for all directions ranged from 0.84 to 0.92, SEM values ranged from 2.01 to 3.10%, SWC values ranged from 3.67 to 5.12%, and MDC95% values ranged from 5.58 to 8.60% (Table 3). For test-retest reliability, the ICC values for all directions ranged from 0.81 to 0.92, SEM values ranged from 1.80 to 2.97%, SWC values ranged from 3.75 to 5.61%, and MDC95% values ranged from 4.98–8.24%(Table 4).

The two lines above and below the mean difference in the Bland Altman plot represent the 95% limits of agreement(95%LOA), indicating the size of the measuring errors. The Bland Altman plots showed that almost

Table 3 Inter-rater reliability for the YBT-LQ in the flexible flatfoot college students

YBT-LQ	ICC (2,1)	95%CI	SEM (%)	SWC (%)	MDC95% (%)
AT	0.84	[0.69,0.92]	2.67	3.99	7.39
PM	0.92	[0.84,0.96]	2.44	5.11	6.75
PL	0.87	[0.74,0.94]	3.10	5.12	8.60
CS	0.89	[0.79,0.92]	2.01	3.67	5.58

ICC: Intra Class Correlation; 95% CI: 95% confidence interval; SEM: Standard Error of Measurement; SWC: Smallest worthwhile change; MDC95%: Minimal detectable change using a 95% confidence interval

Table 4 Test-retest reliability for the YBT-LQ in the flexible flatfoot college students

YBT-LQ	ICC (3,1)	95%CI	SEM (%)	SWC (%)	MDC95% (%)
AT	0.81	[0.63,0.90]	2.76	3.78	7.64
PM	0.90	[0.80,0.95]	2.97	5.61	8.24
PL	0.89	[0.79,0.95]	2.92	5.32	8.08
CS	0.92	[0.83,0.96]	1.80	3.75	4.98

ICC: Intra Class Correlation; 95% CI: 95% confidence interval; SEM: Standard Error of Measurement; SWC: Smallest worthwhile change; MDC95%: Minimal detectable change using a 95% confidence interval

all points are within the 95%LOA. For each of the three test directions and the composite score, the difference between the two raters and test-retest measurements plotted against the mean difference was close to zero (Figs. 1 and 2).

Discussion

To the best of our knowledge, this is the first study to examine the reliability of the YBT-LQ in a population of college students with flexible flatfoot. This study provides preliminary evidence of the reliability of the YBT for use with a population of college students with flexible flatfoot, and the results of the study indicate that the YBT-LQ has "good" to "excellent" inter-rater and test-retest reliability. The inter-rater point estimates for ICCs ranged from "moderate" to "excellent" on the anterior (Table 3, first line, [0.69–0.92]) and posterolateral (Table 3, third line, [0.74–0.94]), ranged from "good" to "excellent" on the posteromedial (Table 3, second line, [0.84–0.96]) and the composite scores (Table 3, fourth line, [0.79–0.92]). The test-retest point estimates for ICCs ranged from "moderate" to "excellent" on the anterior (Table 4, first
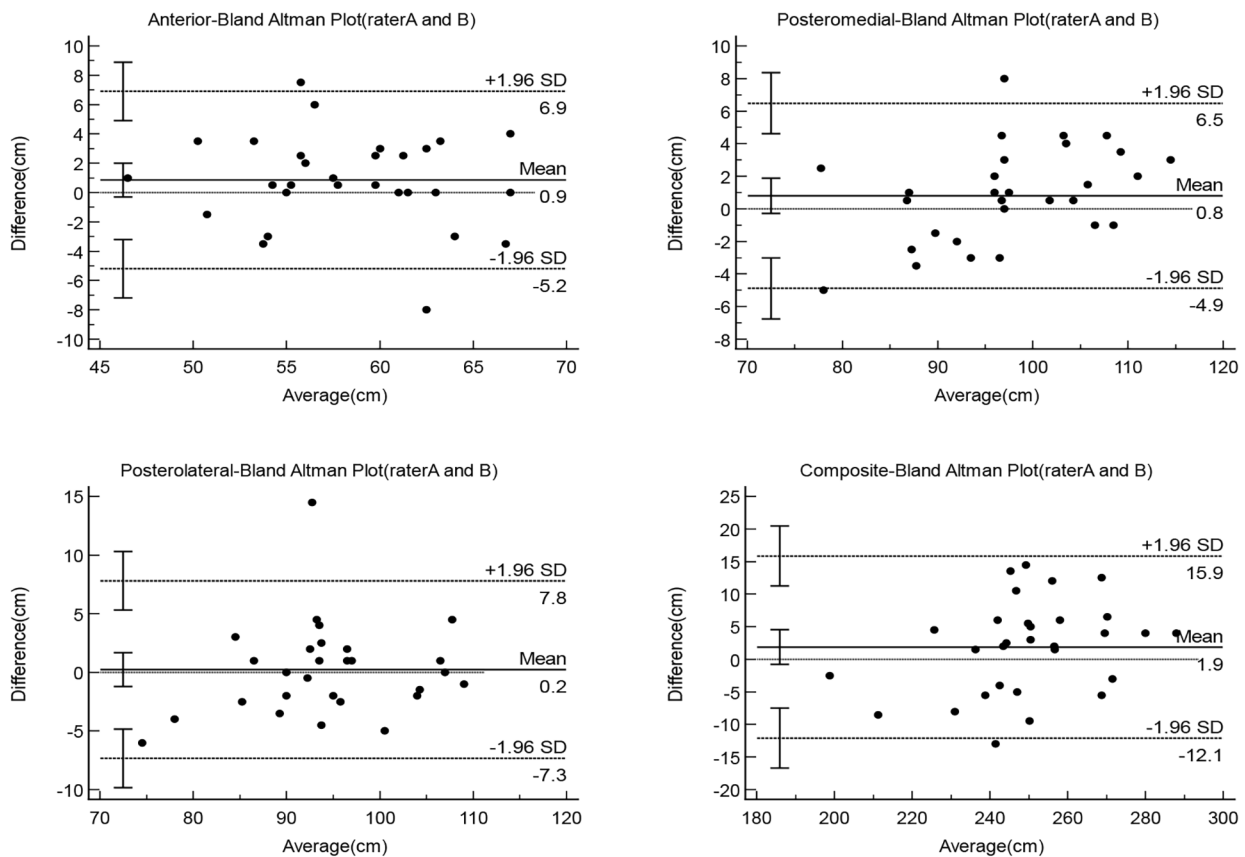

Fig. 1 Bland-Altman Plots with 95% limits of agreement illustrate measurement errors against true values by plotting each of the three YBT-LQ directions (anterior, posteromedial, and posterolateral) and the composite score differences between the two raters. Plot of the mean (x-axis) of the measurements of two raters against the difference (y-axis) between two raters
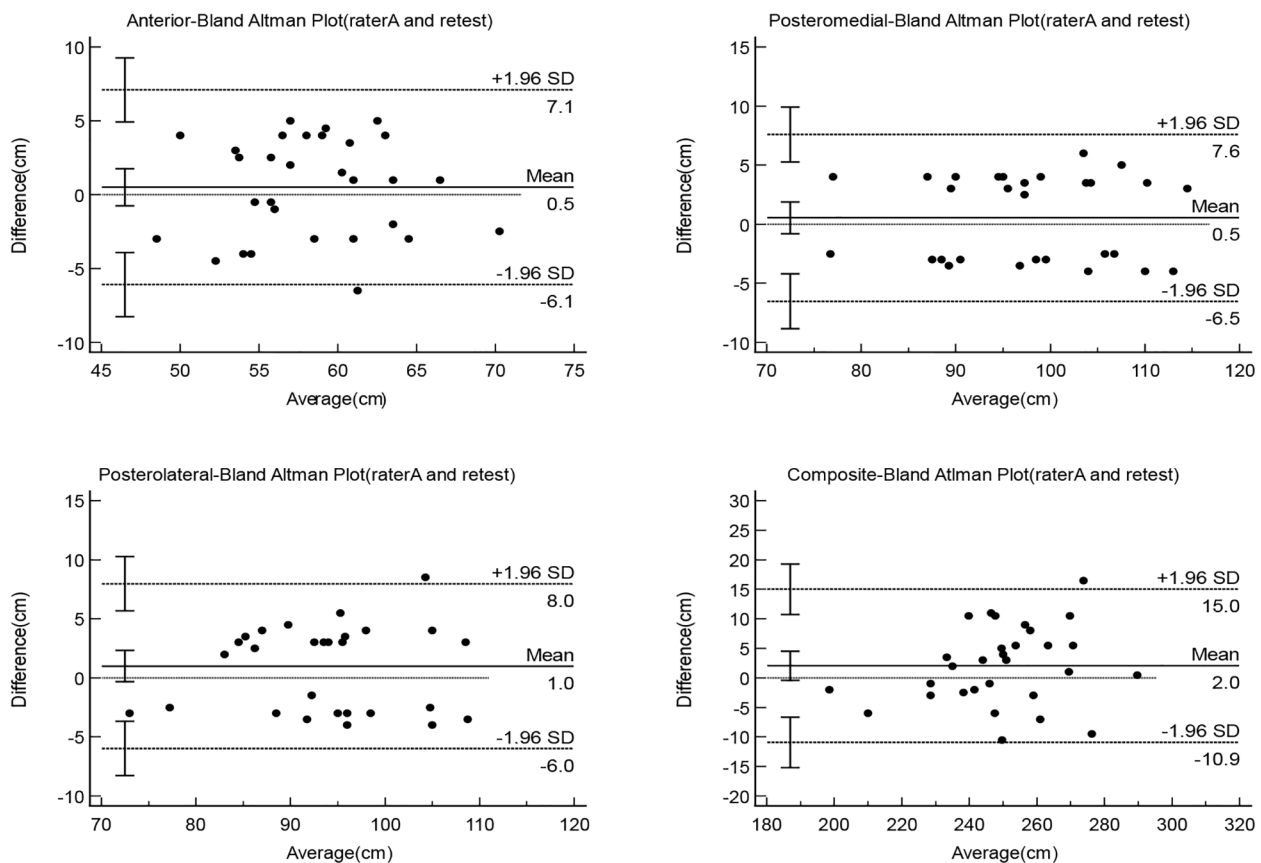
Zheng *et al. BMC Sports Science, Medicine and Rehabilitation* (2024) 16:40

Page 6 of 8

Fig. 2 Bland-Altman Plots with 95% limits of agreement illustrating measurement errors against true values by plotting each of the three YBT-LQ directions (anterior, posteromedial, and posterolateral) and the composite score differences between the rater A test and retest. Plot of the mean (x-axis) of the measurements of test and retest against the difference (y-axis) between test and retest

line, [0.63–0.90]), ranged from "good" to "excellent" on the posteromedial (Table 4, second line, [0.80–0.95]), posterolateral (Table 4, third line, [0.79–0.95]) and the composite scores (Table 3, fourth line, [0.83– 0.96]).. Moderate to excellent inter-rater, test-retest reliability has been reported in previous research among 51 health populations aged $19\sim50$ years old (ICC$0.79\sim0.86$, SEM2%~4%) [28]; 178 teenagers aged $11\sim19$ years old(ICC$0.4\sim0.96$, SEM1.77%~5.81%) [35]; 110 high school athlete(ICC$0.63\sim0.89$, SEM1.94%~4.17%) [36]. These results support the findings of our study and import the YBT-LQ as a valid test method for detecting changes in dynamic equilibrium over time at the group level.

Absolute reliability is the degree of variability of an individual's repeated measurements. The SEM, SWC, and MDC 95% were calculated to quantify the error due to repeated measurements. Lower SEM and MDC 95% suggest strong reliability of the measurements [33]. It can be said that the measurement error is minimal and the measurement is reliable when the SEM value is less than 10% of the highest or average measurement value [37]. The

inter-rater and test-retest have a rather smaller SEM of the YBT-LQ between 1.80 and 3.10%. SWC ranged from 4.98 to 8.60%, all greater than SEM, indicating that the Y balance test can detect smallest worthwhile changes. Further MDC95%, which represents the change needed to identify clinically relevant effects between repeated measures ranged from 4.98 to 8.60%. Compared to other studies, our MDC95% values are close to those(5%~11%) reported by Foldager et al. [28]. The difference between our values and those of Schwiertz et al. [36] might be due to the investigated cohort. Our study was in a college student's cohort, while Schwiertz investigated 178 adolescents in grades 6–11, who may have shown the worse ability to understand and perform YBT compared to college students. Notably, the composite score showed the most reliable results with the higher ICC and the lowest SEM and MDC95% for both inter-rater and test-retest. Therefore, it is recommended that the composite score of the YBT be used in screening to quickly assess dynamic balance in a college student population with flexible flatfoot.

However, this study has some limitations. Because raters were involved in the research design process, the Hawthorne effect (i.e., raters were unaware that their judgments would be compared with those of other assessors) could not be avoided and the raters' behavior may be altered as a result of the awareness of being observed. Moreover, for the test and retest reliability test, we were unable to blind the rater to the fact that he was going to take the YBT-LQ test. However, we believe the potential for a recall bias was held to a minimum due to several factors: (1) Hiding test data immediately after completion of the test so that the rater does not have access to each subject's previous score; (2) the data were not analyzed until the end of the study; and (3) the testing and retesting intervals 14 days, with as many as 12 participants being tested on a single day. Given these factors, it is unlikely that the rater had a clear recall of any individual's last test performance. Additionally, the findings of the study revealed that the posterior lateral side had a higher inter-rater measurement error (SEM3.10%) than did the other two directions (SEM2.44~2.67%), which may be since only the lower limb on the side of the flatfoot was tested in this study, whereas the YBT protocol recommends switching between the right and left leg in each test direction. Prolonged single-leg support may increase the possibility of muscle fatigue, which would decrease the dynamic balance [38]. The intervals between each testing direction could be appropriately lengthened for subjects in subsequent investigations to reduce this measurement error. Finally, although previous research protocols have shown that six practice sessions in each direction before the start of the formal test can minimize the learning effect, a paired-samples t-test would have made this experimental protocol more rigorous. We hope that future researchers will improve on the experimental protocol of this study by fully recording the values of the three valid experiments and conducting a paired-samples t-test (trial-by-trial values) to explore whether six practice sessions to minimize the learning effect is reliable and provide more reliable evidence for the application of the YBT-LQ.

In summary, this study demonstrates that the YBT-LQ has "good" to "excellent" intra-rater and inter-rater reliability as a simple, rapid method of assessing dynamic balance ability in a college student flatfoot population. It can be applied to screen this population in college so that targeted rehabilitation training and education in advance can be provided to avoid sports injuries and improve the physical fitness of college students. However, distance-attainment criteria have not yet been developed in various populations. Future research could focus on the reliability of the YBT and the validity of the YBT in predicting sports injuries using standardized test protocols in different clinical settings and populations. In addition, different distance-to-reach standards could be developed for different groups through standardized data collection and prospective studies.

Conclusions

The inter-rater and test-retest reliability of YBT-LQ ranges from "good" to "excellent" as a reliable tool for assessing the dynamic balance ability of college students with flexible flatfoot populations.

Abbreviations

YBT-LQ	Lower Quarter Y Balance Test
AT	Anterior
PM	Posteromedial
PL	Posterolateral
CS	Composite score
BMI	Body mass index
NDT	The navicular drop test
ICC	Intra Class Correlation
SEM	Standard Error of Measurement
SWC	Smallest worthwhile change
MDC95%	Minimal detectable change using a 95% confidence interval

Author contributions
ZHENG YL and FENG RZ designed the research question. ZHENG YL and FENG RZ planned and performed the Y balance test and data recording. HU WY and HUANG P managed and analyzed the data. ZHENG YL, FENG RZ, and HU WY wrote the main parts of the manuscript. All authors contributed to the critical review of draft manuscripts and approved the final manuscript.

Data availability
The data generated and analyzed during the present study are not publicly available due to ethical restrictions but are available from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate
Ethical permission was given by the Human Ethics Committee at Beijing Sports University. Approval ID: 2023154 H. Approval Date: 05/09/2023. Participants written informed consent was obtained prior to the start of the study, and procedures were conducted according to the Declaration of Helsinki.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

References
1. Abd-Elmonem AM, El-Negamy EH, Mahran MA, Ramadan AT. Clinical and radiological outcomes of corrective exercises and neuromuscular electrical

stimulation in children with flexible flatfeet: a randomized controlled trial. Gait Posture. 2021;88:297–303.

2. Moon D, Jung J. Effect of incorporating short-foot exercises in the Balance Rehabilitation of flat foot: a Randomized Controlled Trial. Healthc (Basel). 2021;9:1358.

3. He H, Liu W, Teraili A, Wang X, Wang C. Correlation between flat foot and patellar instability in adolescents and analysis of related risk factors. J Orthop Surg (Hong Kong). 2023;31:10225536231171057.

4. Kim E-K, Kim JS. The effects of short foot exercises and arch support insoles on improvement in the medial longitudinal arch and dynamic balance of flexible flatfoot patients. J Phys Ther Sci. 2016;28:3136–9.

5. Şahin FN, Ceylan L, Küçük H, Ceylan T, Arıkan G, Yiğit S, et al. Examining the relationship between Pes Planus Degree, Balance and Jump performances in athletes. Int J Environ Res Public Health. 2022;19:11602.

6. Takabayashi T, Edama M, Inai T, Kubo M. Differences in rearfoot, midfoot, and forefoot kinematics of normal foot and flatfoot during running. J Orthop Res. 2021;39:565–71.

7. Gonell AC, Romero JAP, Soler LM. Relationship between the Y balance test scores and soft tissue injury incidence in a soccer team. Int J Sports Phys Ther. 2015;10:955–66.

8. Hartley EM, Hoch MC, Boling MC. Y-balance test performance and BMI are associated with ankle sprain injury in collegiate male athletes. J Sci Med Sport. 2018;21:676–80.

9. Gkrilias P, Zavvos A, Fousekis K, Billis E, Matzaroglou C, Tsepis E. Dynamic balance asymmetries in pre-season injury-prevention screening in healthy young soccer players using the modified Star Excursion Balance Test-a pilot study. J Phys Ther Sci. 2018;30:1141–4.

10. Barrera-Domínguez FJ, Carmona-Gómez A, Tornero-Quiñones I, Sáez-Padilla J, Sierra-Robles Á, Molina-López J. Influence of dynamic balance on jumping-based asymmetries in Team Sport: A between-sports comparison in basketball and Handball athletes. Int J Environ Res Public Health. 2021;18:1866.

11. Jaber H, Lohman E, Daher N, Bains G, Nagaraj A, Mayekar P, et al. Neuromuscular control of ankle and hip during performance of the star excursion balance test in subjects with and without chronic ankle instability. PLoS ONE. 2018;13:e0201479.

12. Ganesh GS, Chhabra D, Mrityunjay K. Efficacy of the star excursion balance test in detecting reach deficits in subjects with chronic low back pain. Physiother Res Int. 2015;20:9–15.

13. Plisky PJ, Gorman PP, Butler RJ, Kiesel KB, Underwood FB, Elkins B. The reliability of an instrumented device for measuring components of the Star Excursion Balance Test.

14. Powden CJ, Dodds TK, Gabriel EH, The reliability of the star excursion balance test, and lower quarter Y-balance test in healthy adults. A systematic review. Intl J Sports Phys Ther. 2019;14:683–94.

15. Wang Z, An G, Zhang W, Yang G. The effect of jazz dance on physical and mental health of students with different physical fitness. J Sports Med Phys Fitness. 2019;59:880–5.

16. Wang J. The association between physical fitness and physical activity among Chinese college students. J Am Coll Health. 2019;67:602–9.

17. Tsai L-C, Yu B, Mercer VS, Gross MT. Comparison of different structural foot types for measures of standing postural control. J Orthop Sports Phys Ther. 2006;36:942–53.

18. Garrison JC, Bothwell JM, Wolf G, Aryal S, Thigpen CA. Y balance test™ anterior reach symmetry at three months is related to single leg functional performance at time of return to sports following anterior cruciate ligament reconstruction. Int J Sports Phys Ther. 2015;10:602–11.

19. Mayer SW, Queen RM, Taylor D, Moorman CT, Toth AP, Garrett WE, et al. Functional testing differences in Anterior Cruciate Ligament Reconstruction patients released Versus not released to return to Sport. Am J Sports Med. 2015;43:1648–55.

20. Boyle MJ, Butler RJ, Queen RM. Functional Movement Competency and Dynamic Balance after Anterior Cruciate Ligament Reconstruction in adolescent patients. J Pediatr Orthop. 2016;36:36–41.

21. Tahmasbi A, Shadmehr A, Attarbashi Moghadam B, Fereydounnia S. Comparison between the effects of tibialis posterior versus fibularis longus Kinesio taping on foot posture, physical performance, and dynamic balance in young women with flexible flatfoot. Sport Sci Health. 2023;19:147–54.

22. Freund JE, Stetts DM, Oostindie A, Shepherd J, Vallabhajosula S. Lower quarter Y-balance test in healthy women 50–79 years old. J Women Aging. 2019;31:475–91.

23. Greenberg ET, Barle M, Glassmann E, Jung M-K. Interrater and test-retest reliability of the y balance test in healthy, early adolescent female athletes. Int J Sports Phys Ther. 2019;14:204–13.

24. Koo TK, Li MY. A Guideline of selecting and reporting Intraclass correlation coefficients for Reliability Research. J Chiropr Med. 2016;15:155–63.

25. Shaffer SW, Teyhen DS, Lorenson CL, Warren RL, Koreerat CM, Straseske CA, et al. Y-balance test: a reliability study involving multiple raters. Mil Med. 2013;178:1264–70.

26. Kottner J, Audigé L, Brorson S, Donner A, Gajewski BJ, Hróbjartsson A, et al. Guidelines for reporting reliability and Agreement studies (GRRAS) were proposed. J Clin Epidemiol. 2011;64:96–106.

27. McPoil TG, Cornwall MW, Medoff L, Vicenzino B, Forsberg K, Hilz D. Arch height change during sit-to-stand: an alternative for the navicular drop test. J Foot Ankle Res. 2008;1:3.

28. Foldager FN, Aslerin S, Mechlenburg I, Interrater. Test-retest Reliability of the Y Balance Test: A Reliability Study Including 51 Healthy Participants. 2023.

29. Sipe CL, Ramey KD, Plisky PP, Taylor JD, Y-Balance Test. A Valid and Reliable Assessment in older adults. J Aging Phys Act. 2019;27:663–9.

30. Plisky PJ, Rauh MJ, Kaminski TW, Underwood FB. Star Excursion Balance Test as a predictor of lower extremity injury in high school basketball players. J Orthop Sports Phys Ther. 2006;36:911–9.

31. Filipa A, Byrnes R, Paterno MV, Myer GD, Hewett TE. Neuromuscular training improves performance on the star excursion balance test in young female athletes. J Orthop Sports Phys Ther. 2010;40:551–8.

32. Weir JP. Quantifying test-retest reliability using the Intraclass correlation coefficient and the SEM. J Strength Cond Res. 2005;19:231.

33. Liaw L-J, Hsieh C-L, Lo S-K, Chen H-M, Lee S, Lin J-H. The relative and absolute reliability of two balance performance measures in chronic stroke patients. Disabil Rehabil. 2008;30:656–61.

34. Marocolo M, Simim MAM, Bernardino A, Monteiro IR, Patterson SD, da Mota GR. Ischemic preconditioning and exercise performance: shedding light through smallest worthwhile change. Eur J Appl Physiol. 2019;119:2123–49.

35. Schwiertz G, Brueckner D, Schedler S, Kiss R, Muehlbauer T. Performance and reliability of the lower quarter Y balance test in healthy adolescents from grade 6 to 11. Gait Posture. 2019;67:142–6.

36. Smith LJ, Creps JR, Bean R, Rodda B, Alsalaheen B. Performance and reliability of the Y-Balance Test™ in high school athletes. J Sports Med Phys Fitness. 2018;58:1671–5.

37. Hyong IH, Kim JH. Test of intrarater and interrater reliability for the star excursion balance test. J Phys Ther Sci. 2014;26:1139–41.

38. Gribble PA, Hertel J, Denegar CR, Buckley WE. The effects of fatigue and chronic ankle instability on dynamic Postural Control. J Athl Train. 2004;39:321–9.

Publisher's Note